



**Assessing synovitis in the hands in patients with rheumatoid arthritis by ultrasound
an agreement study exploring the most inflammatory active side from two Norwegian
trials**

Terslev, Lene; Christensen, Robin; Aga, Anna Birgitte; Sexton, Joe; Haavardsholm, Espen A.;
Hammer, Hilde B.

Published in:
Arthritis Research & Therapy

DOI:
[10.1186/s13075-019-1930-y](https://doi.org/10.1186/s13075-019-1930-y)

Publication date:
2019

Document version
Publisher's PDF, also known as Version of record

Document license:
[CC BY](#)

Citation for published version (APA):
Terslev, L., Christensen, R., Aga, A. B., Sexton, J., Haavardsholm, E. A., & Hammer, H. B. (2019). Assessing synovitis in the hands in patients with rheumatoid arthritis by ultrasound: an agreement study exploring the most inflammatory active side from two Norwegian trials. *Arthritis Research & Therapy*, 21, [166].
<https://doi.org/10.1186/s13075-019-1930-y>

RESEARCH ARTICLE

Open Access



Assessing synovitis in the hands in patients with rheumatoid arthritis by ultrasound: an agreement study exploring the most inflammatory active side from two Norwegian trials

Lene Terslev^{1,2*} , Robin Christensen^{2,3}, Anna-Birgitte Aga⁴, Joe Sexton⁴, Espen A. Haavardsholm^{4,5} and Hilde B. Hammer⁴

Abstract

Objective: To assess if the right hand, the dominant hand, or the hand with more clinically swollen joints (SwJ) is per se the most inflamed and exhibits the greatest change during treatment and hence preferred for unilateral scoring of synovitis by ultrasound in rheumatoid arthritis (RA) patients.

Methods: Using data from two previously published Norwegian RA patient cohorts initiating treatment, bilateral metacarpophalangeal joint 1–5, proximal phalangeal joint 2+3, and wrists were evaluated by ultrasound. Using a 0–3 scoring system a grey-scale (GS), power Doppler (PD) and global synovitis score (GLOESS) was calculated for each hand (0–30). For precision, a difference of $< \pm 3$ in sum score was pre-specified as indicating clinically insignificant difference in inflammatory activity for all three scores.

Results: Four hundred thirty-seven RA patients were included. Baseline ultrasound inflammation was statistically significantly higher in hands with more vs fewer SwJ ([mean difference, 95%CI] GS sum score 2.21[1.30 to 3.12], PD sum score 1.70 [0.94 to 2.47] and GLOESS 2.31[1.36 to 3.26]) and also exhibited significantly more change for all sum scores at 3 months follow-up (GS sum score 1.34 [0.60 to 2.08], PD sum score 1.17 [0.44 to 1.91], and GLOESS 1.43 [0.63 to 2.22]). No such differences were found between the dominant and the non-dominant or the right and the left hands at any time points.

Conclusion: The hand with clinically more SwJ is statistically more inflammatory active according to GS, Doppler, and GLOESS sum scores, exhibits a change during treatment, and is potentially the best choice for unilateral scoring systems.

Keywords: Synovitis, Rheumatoid arthritis, Ultrasound, Doppler, Inflammation, Dominant, Hand

* Correspondence: terslev@dadlnet.dk

¹Center for Rheumatology and Spine Diseases, Rigshospitalet, Valdemar Hansens vej 17, Glostrup, DK-2600 Copenhagen, Denmark

²Musculoskeletal Statistics Unit, The Parker Institute, Bispebjerg and Frederiksberg Hospital, Frederiksberg, Denmark

Full list of author information is available at the end of the article



© The Author(s). 2019 **Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated.

Background

Ultrasound has been validated as an outcome measurement tool for assessing synovitis by grey-scale (GS) and Doppler in rheumatoid arthritis (RA) [1, 2]. Though recent publications have indicated that ultrasound examination of all RA patients—if offered a very tight clinical control—may not be necessary for obtaining clinical remission [3, 4], there are several situations in the clinic where ultrasound has a role in monitoring synovitis as an indicator of disease activity. Furthermore, ultrasound is used in clinical trials for assessing treatment response and remission. The components defining synovitis (synovial hypertrophy and hyperemia) are usually scored separately using a semi-quantitative scoring system (0–3) for indicating the grade of severity of the individual synovitis component [5]. In the recently published OMERACT-EULAR combined scoring system [1], it is suggested to apply the highest score of the two components as the final score for the joint. Scoring GS synovial hypertrophy and Doppler activity separately or in combination is sensitive to change during treatment both on joint and patient level—for the latter using a sum score [6]. When applied in a clinical trial context, several reduced joint sets, ranging from 6 to 12 joints, have been proposed for scoring synovitis over the last years [7–11]. These reduced joint sets evaluate synovitis either unilaterally or bilaterally, aiming at maintaining as much information about the inflammatory load of the patient as seen in the more elaborate and time-consuming joint evaluations (32–78 joints) [12, 13].

Scoring synovitis unilaterally will by far reduce the examination time [9], and the unilateral 7-joint count by Backhaus et al. [7] is often used in trials. In unilateral scoring systems, the suggestion is to choose either the dominant side [7, 8] or the clinically most affected side defined as the hand with the most tender and/or swollen joints [9, 14, 15]. The dominant side is also commonly the chosen side in magnetic resonance imaging (MRI), where only one side can be evaluated at a time and is based on data showing that the dominant hand has more erosions on X-ray than the non-dominant hand both at baseline and follow-up [16, 17]. However, there is no evidence per se to support the dominant hand as being more inflammation prone than the non-dominant hand at time of treatment initiation. Because of the lack of evidence-based guidance, we decided to assess which selection is likely to identify the more inflamed hand, a research-on-research project that could provide information about RA presentation of inflammation and could potentially serve as a guide for the “default option” in unilateral scoring systems among ultrasonographers in rheumatology and thus ensure better homogeneity among future studies and clinical reports.

The primary aim of this study was to investigate if one hand is always more inflamed than its counterpart as this would have implications for a unilateral scoring system applied in the clinic and for the application in clinical trials. This was assessed by evaluating if the right hand, the dominant hand (as indicated by MRI), or the hand with clinically more swollen joints is more likely to have higher ultrasound measured inflammation than its counterpart at the time of inclusion, as well as for changes during treatment from baseline to 3 months follow-up.

Patients and methods

This research-on-research project was performed according to a pre-defined Statistical Analysis Plan that was uploaded in advance (see Additional file 1). The study was designed to identify at baseline the hand with most ultrasound-verified inflammation and hence most appropriate for potential unilateral ultrasound assessment, and further to investigate the change in inflammation of this selection at 3 months follow-up. Hence, the follow-up study was not about assessing treatment effect over time, but about evaluating the ability of the inflammation to change.

This study is a result of secondary analyses based on two independent ultrasound datasets including an early RA cohort (ARCTIC trial; [ClinicalTrials.gov](https://clinicaltrials.gov) database [NCT01581294]) and an established RA cohort (ULRABIT trial; [Anzctr.org.au](https://anzctr.org.au) database [ACTRN12610000284066]) initiating conventional synthetic disease-modifying antirheumatic drugs (DMARDs) in the early cohort and initiating or switching biological DMARD (bDMARD) in the established cohort. From the ARCTIC trial (recruited between September 2010 and April 2013), 238 patients with early RA were included, and in the ULRABIT trial (recruited from January 2010 to June 2013), 212 patients with established RA were included.

Ultrasonography assessments

In these analyses, the hands were used as model evaluating metacarpophalangeal joint (MCP) 1–5, proximal interphalangeal joint (PIP) 2 and 3, and the wrist. The selection of PIP joints was based on the study of Backhaus et al. [7]. In the original ULRABIT trial, extensive ultrasound examinations were performed by a single experienced sonographer, and in the ARCTIC trial, by several experienced sonographers who had been trained in annual ultrasound workshops with both static and dynamic hands-on exercises to calibrate readers. In both studies, synovitis was defined hypoechoic synovial hypertrophy with or without Doppler activity including all hypoechogenic tissue inside the joint capsule as synovitis. Though effusion may be hypoechoic, it is mostly anechoic and a rare finding in RA wrists and finger

joints in contrast to knees and MTP joints and therefore not likely to influence the current analysis. Hence for the hands, the definition for synovitis is in line with the modified OMERACT definition [1].

A 0–3 semiquantitative scoring system for both GS and power Doppler findings was applied in each of the following 36 joints: MCP 1–5, PIP 2–3, radiocarpal, midcarpal, distal radioulnar, elbow, knee, talocrural, and metatarsophalangeal (MTP) 1–5 bilaterally using the ultrasound atlas by Hammer et al. as a reference which had demonstrated a high inter- and intra-observer reliability [18]. In the ULRABIT study, a Siemens Antares Sonoline machine was used (Siemens Medical solutions, Mountain view, CA, USA) with a linear probe (5–13 MHz and setting at 11.4 MHz) and unchanged Doppler settings optimized for slow flow [19]. In the ARTIC study, the same Siemens Antares Sonoline machines or GE Logiq E9's (GE Medical Systems Ultrasound and Primary Care Diagnostics, Wauwatosa, WI, USA) were used at baseline and follow-up in the participating centers (11 hospitals), all with linear probes and identical unchanged Doppler settings optimized for slow flow [19].

In the current analyses, both GS and power Doppler sum scores, as well as the OMERACT-EULAR combined score (GLOESS) [1], were calculated for each hand (for all sum scores, the range is 0–30). The highest score of GS and PD was defined as the GLOESS score for the individual joint.

A difference of less than 3 in ultrasound sum scores between the hands was perceived by the authors to be clinically insignificant based on expert opinion and was prespecified in the statistical analysis plan (Additional file 1).

Laboratory and clinical examinations

For both trials, each visit included laboratory assessments of erythrocyte sedimentation rate (ESR) and C-reactive protein (CRP, mg/L) and 0–100-mm visual analogue scales (VAS) for assessor's and patient's global assessments of disease activity. In the ARTIC study, anticyclic citrullinated peptide (anti-CCP) and rheumatoid factor (RF) was assessed at baseline whereas this had been assessed prior to the study for the ULRABIT cohort. For the early RA cohort, 44 swollen joint counts (44 SJC) and Ritchie Articular Index were performed, while 28 swollen and tender joint counts (28 SJC and 28 TJC, respectively) were performed in the established RA cohort. The Disease Activity Score (DAS) was calculated in the early RA cohort, whereas the DAS28 was calculated in the patients with established RA; both scores were based on ESR.

Statistical methods

In this agreement study, our objective was to see whether various approaches showed sufficient agreement

to be applied interchangeably. Our initial approach was based on the Bland-Altman graphical techniques [20] to visually present a potential pattern between left versus right hand as the plot of difference against the mean of the two. This step also included a plot of the data and draw the line of equality on which all points would lie if the two ultrasound measures (left versus right) gave exactly the same reading every time. This scatter plot helps the eye in gauging the degree of agreement between measurements. Also, we use a plot of difference (right versus left) against mean which allows us to investigate any possible relationship between the measurement error and the true value. By calculating mean difference (Δ : left – right) and the standard deviation of the differences ($SD\Delta$), we would expect most of the differences to lie between $\Delta \pm 1.96 \times SD\Delta$ if the differences are normally distributed where 95% of differences will lie between these limits [20].

According to our prespecified statistical analysis plan (Additional file 1), we wanted to evaluate whether there is a potential difference in ultrasound inflammation or two contextual factors: dominant hand and the hand with most swollen joints, respectively. Using SAS for Mixed Models (PROC MIXED), the similarity between measures was tested. The mixed effect model took into account any differences in patient characteristics (e.g., disease duration) between groups by including a fixed factor for cohort (i.e., adjusting for the “cohort effect”). The least squares mean values and the difference between them are reported based on a statistical model including a factor for the specific analysis (i.e., dominant/right/swollen), and trial (ARCTIC and ULRABIT, respectively) as fixed effects, with the patient-ID as a random effect. We a priori defined a reasonable equivalence margin between measures, to be a 95% confidence interval around the observed paired mean difference: –2.99 to +2.99 (Additional file 1).

Results

Baseline and follow-up demographics for the combined cohort as well as for the sub-cohorts are shown in Table 1. The total cohort consisted of 442 patients and comprised the ARCTIC cohort with 230 DMARD naïve patients with early RA with indication for methotrexate (as 8 of the original 238 patients were subsequently excluded from the study as they did not receive treatment) and the ULRABIT cohort with 212 established RA patients with indication for bDMARDs.

Of the 442 patients, only 437 (99%) were included in the ultrasound evaluation as five patients from the ARTIC cohort were enrolled without an ultrasound examination at baseline whereas all 212 patients in the ULRABIT cohort had an ultrasound examination at time of inclusion. At 3 months follow-up, 295 patients had an

ultrasound examination—116 patients from the ARTIC study and 179 patients from the ULRABIT study. For the whole cohort, 21 (5%) patients had left hand dominance and 410 (95%) had right hand dominance, and for 11 patients, the information was missing. Similarly, for the whole cohort, 183 (42%) patients had the right hand as the most clinically affected hand and 131 (30%) had the left hand as the most affected hand at time of treatment initiation, and for 123 (28%) patients, both hands were clinically equally affected (same number of swollen joints in both hands), and for one patient, the clinical joint information was missing.

Differences in inflammatory activity between hands

To assess if any systematic errors could potentially be found between the right- and left-hand side, Bland-Altman scatter plots were made (Fig. 1) to give a visual impression of potential differences and to assess if further analysis should be carried out. Though there seemed to be an even distribution between the right and the left hand for the GS sum score, Doppler sum score, and the GLOESS score, the right hand appeared to be potentially more involved for all three sum scores than the left.

Subsequent analyses were carried out to assess if the potential differences were present between the dominant and non-dominant hands, the right and the left hands, and the more clinically affected hand and the less clinically affected. The analyses of differences in sum scores between the dominant versus the non-dominant hand revealed that the dominant hand was *not* more inflammatory active than the non-dominant (mean difference, 95%CI) for GS sum score -0.69 (-1.61 to 0.24 ; $p = 0.15$; both confidence limits within the equivalence margin), power Doppler sum score -0.58 (-1.35 to 0.20 ; $p = 0.14$; both confidence limits within the equivalence margin), and GLOESS -0.79 (-1.76 to 0.18 ; $p = 0.11$; both confidence limits within the equivalence margin) (Table 2). Similar results were found for the right versus left hand where the mean difference for the GS sum score was -0.75 (-1.68 to 0.18 ; $p = 0.11$; both confidence limits within the equivalence margin), for the power Doppler sum score -0.61 (-1.38 to 0.16 ; $p = 0.12$; both confidence limits within the equivalence margin) and for the GLOESS -0.89 (-1.85 to 0.08 ; $p = 0.07$; both confidence limits within the equivalence margin).

When analyzing the importance of SwJ in the 314 of the 437 patients having differences in SwJ between the two sides, the analyses of the hand with more swollen joints showed a statistically significant difference in inflammatory activity as compared to the hand with less swollen joints for all three US variables: GS sum score 2.21 (1.30 to 3.12 ; $p < .0001$), Doppler sum score 1.70 (0.94 to 2.47 ; $p < .0001$), and GLOESS 2.31 (1.36 to 3.26 ;

$p < .0001$). However, with the pre-specified clinically important difference of 3 in sum scores, only the GS and GLOESS sum score may have a clinical impact whereas only a borderline clinical significance was found for the Doppler sum score as the upper confidence limit was lower than the equivalence margin of 3.

Differences in inflammatory activity assessed over time

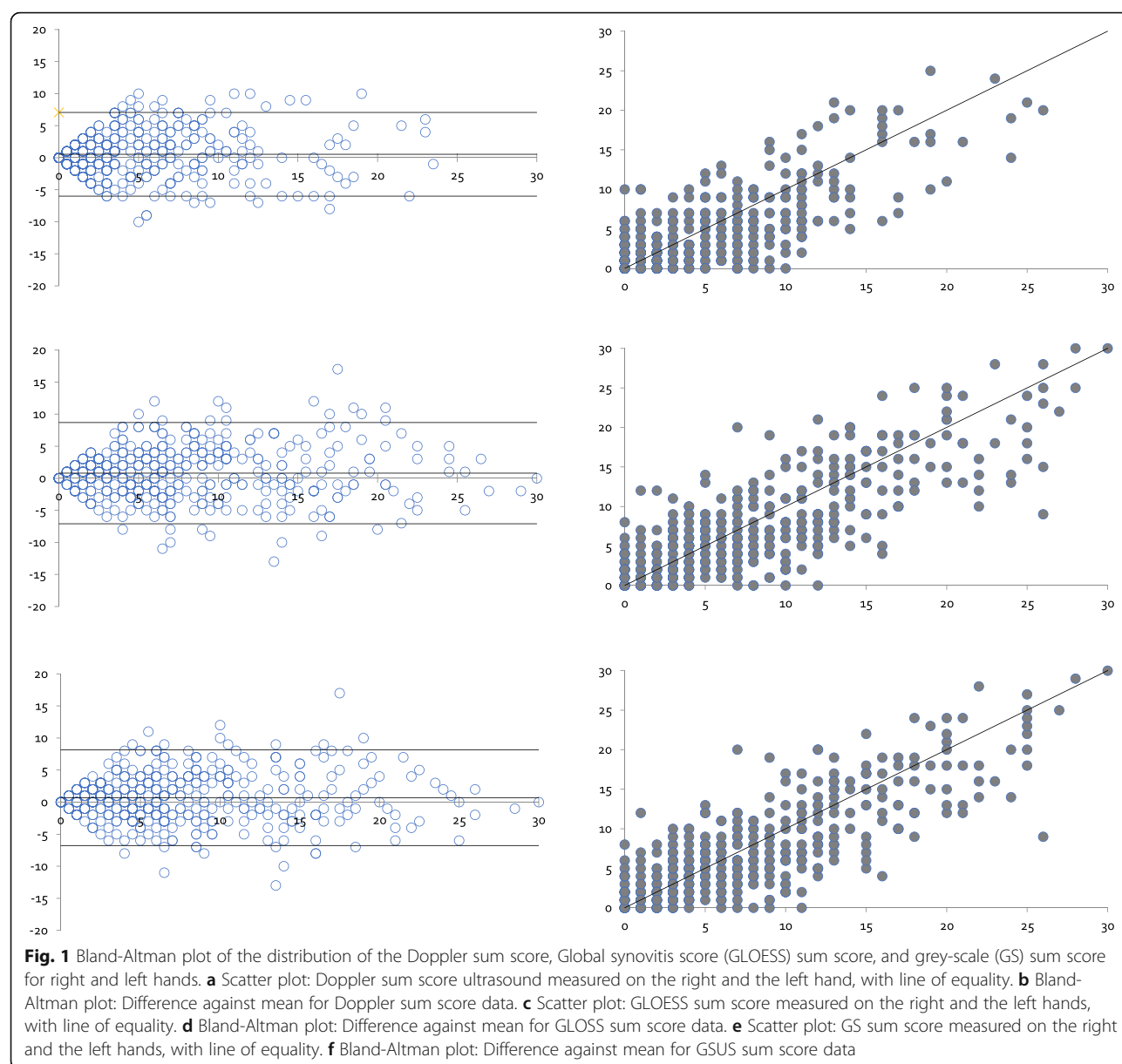
Clinically, there was a significant change in DAS for the ARCTIC cohort and in DAS28 (ESR) for the ULRABIT cohort from baseline to 3 months (for ARCTIC (mean (SD)) -1.75 (0.94) and for ULRABIT -1.08 (1.26) – $p < 0.001$ for both) [4].

Differences in (exhibiting a) change during treatment were assessed for GS sum score, Doppler sum score, and GLOESS for the dominant versus the non-dominant hand, the right versus left hand, and the clinically more affected hand versus the less clinically affected hand (Table 2) and found that the clinically more affected hand changed more than the less clinically affected hand (mean difference, 95%CI) for GS sum score 1.34 (0.60 to 2.08 ; $p < 0.0004$), Doppler sum score 1.17 (0.44 to 1.91 , $p < 0.002$) and GLOESS 1.43 (0.63 to 2.22 ; $p < 0.0005$). No differences in change over time were found for the right versus left hand and the dominant versus the non-dominant hand.

Discussion

In this agreement study comprising both early and established RA patients, we assessed differences in inflammatory activity in the right hand, the dominant hand, and the hand with clinically more swollen joints as compared to its counterpart. We found that the clinically more affected hand at baseline (determined as the hand with more clinically swollen joints) was also the more inflammatory active side—independent of using GS sum score, Doppler sum score, or the GLOESS. Similarly, the clinically most affected side was also found to be the hand exhibiting a greater change for all sum scores during treatment. These findings are to some extent in line with the subset of studies applying the unilateral 7-joint score where the most affected hand evaluated is the hand with the most swollen and/or tender joint—in these studies, this hand was sensitive to change after onset or switch of therapy (DMARDs and/or biologic) [10, 11]. In the current study, we did not evaluate the differences in inflammatory activity between hands with more and less tender joints as tender joints in contrast to swollen joints are not necessarily related to inflammatory joint involvement [21–24].

In clinical trials and in a daily clinical setting when assessing a treatment effect, it is common practice to choose the most inflamed joints for evaluation as they are more likely to exhibit a change during treatment. It



was therefore of interest to evaluate if for unilateral assessment one hand was by nature always more inflamed than its counterpart based solely, i.e., on handedness as this would have impact on clinical trials and for standardized assessments in the clinic. Based on the findings in this research-on-research study applying a clinically significant cutoff of 3 in ultrasound sum scores, a default decision cannot be made on which side to scan. A clinical joint evaluation for SwJ is mandatory for choosing the optimal side for assessing treatment effect. However, the predictive value of ultrasound for treatment response and for obtaining remission has yet to be determined. Our data may however be used in subsequent prognostic factor research to explore whether the most involved hand is more appropriate (than any arbitrary decision)

when trying to predict outcomes in longitudinal cohort studies and to predict who will benefit from new potential effective drugs.

As the dominant hand is often chosen in MRI studies based on reports that this hand is having more erosive changes over time [16, 17], one could speculate that this would be linked to the dominant hand being more inflammatory active than the non-dominant hand. In ultrasound studies, higher levels of inflammation in RA patients with both active disease and in remission have been shown to be related to erosive progression [25–27]; however, we were not able to show that the dominant hand was significantly more inflammatory active than the non-dominant hand nor was there any difference per se in inflammatory activity between the right and left

Table 2 Ultrasound inflammatory activity in the dominant hand versus the non-dominant hand, in the hand with clinically more swollen joint versus less swollen joints, and in the right hand versus left hand using different composite scores for the early and established RA cohorts combined

Baseline				Change values from baseline to 3 months		
	GS sum score (0–30)	Doppler sum score (0–30)	GLOESS (0–30)	Δ GS sum score (0–30)	Δ Doppler sum score (0–30)	Δ GLOESS (0–30)
	Mean (95% CI)	Mean (95% CI)	Mean (95% CI)	Mean (95% CI)	Mean (95% CI)	Mean (95% CI)
*Dominant hand (n = 431)						
Dominant	7.75 (7.09 to 8.40)	4.81 (4.26 to 5.35)	8.07 (7.39 to 8.75)	– 3.11 (– 3.65 to – 2.57)	– 2.57 (– 3.10 to – 2.04)	– 3.40 (– 3.98 to – 2.82)
Non-dominant	7.06 (6.41 to 7.72)	4.22 (3.68 to 4.78)	7.28 (6.60 to 7.96)	– 3.12 (– 3.66 to – 2.58)	– 2.35 (– 2.88 to – 1.82)	– 3.29 (– 3.87 to – 2.71)
Difference	– 0.69 (– 1.61 to 0.24)	– 0.58 (– 1.35 to 0.20)	– 0.79 (– 1.76 to 0.18)	– 0.005 (– 0.76 to 0.75)	0.22 (– 0.52 to 0.96)	0.11 (– 0.70 to 0.92)
P value	0.15	0.14	0.11	0.99	0.57	0.79
*Clinical hand (n = 314)						
Worst clinical (more SJC)	8.51 (7.87 to 9.16)	5.37 (4.83 to 5.91)	8.83 (8.16 to 9.50)	– 3.79 (– 4.32 to – 3.25)	– 3.05 (– 3.57 to – 2.52)	– 4.06 (– 4.63 to – 3.49)
Least clinical (fewer SJC)	6.30 (5.65 to 6.94)	3.66 (3.13 to 4.20)	6.52 (5.85 to 7.19)	– 2.45 (– 2.98 to – 1.92)	– 1.87 (– 2.40 to – 1.35)	– 2.64 (– 3.21 to – 2.06)
Difference	– 2.21 (– 3.12 to – 1.30)	– 1.70 (– 2.47 to – 0.94)	– 2.31 (– 3.26 to – 1.36)	1.34 (0.60 to 2.08)	1.17 (0.44 to 1.91)	1.43 (0.63 to 2.22)
P value	< .0001	< .0001	< .0001	0.0004	0.002	0.0005
*Hand side (n = 437)						
Right side	7.78 (7.12 to 8.44)	4.82 (4.28 to 5.37)	8.12 (7.44 to 8.80)	– 3.19 (– 3.73 to – 2.65)	– 2.64 (– 3.17 to – 2.11)	– 3.49 (– 4.07 to – 2.91)
Left side	7.03 (6.37 to 7.68)	4.21 (3.6 to 4.76)	7.23 (6.55 to 7.91)	– 3.05 (– 3.59 to – 2.51)	– 2.28 (– 2.81 to – 1.75)	– 3.21 (– 3.78 to – 2.63)
Difference	– 0.75 (– 1.68 to 0.18)	– 0.61 (– 1.38 to 0.16)	– 0.89 (– 1.85 to 0.08)	0.14 (– 0.61 to 0.89)	0.36 (– 0.38 to 1.10)	0.29 (– 0.52 to 1.09)
P value	0.11	0.12	0.07	0.71	0.34	0.49

*Analyzed using a factor for the specific analyses (dominant, clinical and hand, respectively), and trial (ARCTIC and ULTRABIT, respectively) as fixed effects, and the patient-ID was applied as a random effect

SJC swollen joint count, Δ delta values from baseline to 3 months, GS grey scale, GLOESS the OMERACT-EULAR combined score

hands. During treatment, neither the dominant hand nor the right hand displayed significantly greater change than the opposite hand.

Reduced joint combinations when using a 78-joint score as gold standard all respond well to biological treatment and all seem to give comparable information about the inflammatory activity in RA patients as a comprehensive ultrasound examination [12]. The best reduced set of joints have yet to be established [13], and though bilateral joint examinations better retain the inflammatory load of the patients than a unilateral evaluation [8], there is an interest in reducing examination time for feasibility reasons in daily clinical practice but also in clinical trials when applying ultrasound as an outcome measurement instrument. One of the often-mentioned reasons for the hampered implementation of ultrasound in the clinic is the lack of time. The current study has used the hands as a model as they are involved in all reduced joint sets [7–11], and examination solely of the hands has previously been shown to detect more than 91% of RA patients in remission with subclinical inflammation [28] and hence appears to retain valuable information about the inflammatory load of the patient.

Though much emphasis has been put on the presence of Doppler activity for assessing the degree of inflammation [3, 4], it has recently been demonstrated that also synovial hypertrophy without Doppler activity may change during treatment [29, 30], and it is therefore interesting that the most affected side for synovitis is independent of applying GS sum score or the GLOESS. The Doppler sum score showed statistically but only borderline clinical significance. This supports that both aspects of the synovitis complex (GS synovial hypertrophy and Doppler activity) are important and that if choosing unilateral scoring systems even when working with an ultrasound equipment with an insensitive Doppler, it is still the hand with the most swollen joints that should be chosen at time of evaluation and for follow-up.

In the current study, we applied a cutoff for clinically important differences in sum score of 3 for synovial hypertrophy, Doppler and GLOESS between the two hands based on expert opinion. One could argue that a lower cutoff would be equally correct, but we chose a stricter approach resulting in GS and GLOESS sum scores showing clinical and statistical differences between the hand with more and less SwJ. The Doppler

sum score, though, was only borderline clinically different, but showed a statistically significant difference. However, the minimal clinically important difference in sum score for GS and Doppler has yet to be established and should be tested in subsequent studies.

The strength of the current study is the sample size and that the analyses include both early and established RA patients making the findings relevant in clinical practice. In addition, the two studies used the similar scoring system based on a US atlas. In the current datasets, 28% of the patients had similar clinical findings in both hands. In such cases, we have no recommendation regarding which side to choose for longitudinal follow-up of a unilateral ultrasound score, and the ultrasonographer may choose freely which side to scan.

Conclusion

No hand is per default more inflamed than its counterpart. The hand with clinically more swollen joints is probably the best choice for monitoring ultrasound activity in clinical practice and in trials and thus the best indicator for the side to choose if a unilateral scoring system is used.

Additional file

Additional file 1: Statistical analysis plan. (PDF 419 kb)

Acknowledgements

The Parker Institute, Bispebjerg and Frederiksberg Hospital, is supported by a core grant from the Oak Foundation (OCAY-13-309). For the UL-RABIT study, we acknowledge our study nurses Anne Katrine Kongtorp and Britt Birketvedt who were important in the organization of the study and performed clinical joint assessments. For the ARCTIC study, we wish to thank the following for the participation in the study: the patients and the ARCTIC investigators Hallvard Fremstad, Tor Magne Madland, Åse Stavland Lexberg, Hilde Haukeland, Erik Rødevand, Christian Høili, Hilde Stray, Anne Lindtner Noraas, Dag Magnar Soldal, and Gunnstein Bakland, as well as all other study personnel.

Authors' contributions

LT and RC performed the post hoc study design; HBH, ABA, and EAH carried out the data acquisition; and RC performed the statistical analyses. LT, RC, ABA, EAH, JS, and HBH carried out data interpretation and manuscript preparation. All authors read and approved the final manuscript.

Funding

The UL-RABIT trial was supported by AbbVie, Pfizer, and Roche in form of study grants to the Department of Rheumatology, DiakonHjmmet Hospital, Oslo, Norway. The ARCTIC trial was supported by Pfizer, UCB, Roche, MSD, and AbbVie.

Availability of data and materials

Please contact the authors for data requests.

Ethics approval and consent to participate

The UL-RABIT and ARCTIC studies were approved by the Regional Committee for Medical and Health Research Ethics South East, Norway, and all patients gave written consent according to the Declaration of Helsinki.

Consent for publication

Not applicable.

Competing interests

Lene Terslev: speaker's fee from Roche, MSD, BMS, Pfizer, AbbVie, Novartis, and Janssen.

Anna-Birgitte Aga: Advisory boards for UCB, AbbVie and Pfizer, honoraria for development of educational material for UCB.

Espen Haavardsholm: Research funding from Pfizer, UCB, Roche, MSD and AbbVie, honoraria as a speaker from Pfizer, UCB, Roche, and AbbVie, honoraria for development of educational material from Pfizer, and has been on advisory board for Pfizer.

Hilde Berner Hammer: Study grants from AbbVie/Pfizer and Roche. Speaker fees from AbbVie, Bristol-Myers Squibb, Roche, UCB Pharma, and Pfizer.

The other authors declare that they have no competing interests.

Author details

¹Center for Rheumatology and Spine Diseases, Rigshospitalet, Valdemar Hansens vej 17, Glostrup, DK-2600 Copenhagen, Denmark. ²Musculoskeletal Statistics Unit, The Parker Institute, Bispebjerg and Frederiksberg Hospital, Frederiksberg, Denmark. ³Department of Clinical Research, Research Unit of Rheumatology, University of Southern Denmark, Odense University Hospital, Odense, Denmark. ⁴Department of Rheumatology, DiakonHjemmet Hospital, Oslo, Norway. ⁵Department of Health Management and Health Economics, University of Oslo, Oslo, Norway.

Received: 7 April 2019 Accepted: 4 June 2019

Published online: 05 July 2019

References

- D'Agostino MA, Terslev L, Aegerter P, Backhaus M, Balint P, Bruyn GA, et al. Scoring ultrasound synovitis in rheumatoid arthritis: a EULAR-OMERACT Ultrasound Taskforce – part 1: definition and development of a standardized, consensus-based scoring system. *RMD Open*. 2017;3:e000428.
- Terslev L, Naredo E, Aegerter P, Wakefield RJ, Backhaus M, Balint P, et al. OMERACT-EULAR Ultrasound Task Force - reliability and applicability of a standardized consensus-based combined synovitis scoring system in rheumatoid arthritis. *RMD Open*. 2017;3:e000427.
- Dale J, Stirling A, Zhang R, Purves D, Foley J, Sambrook M, et al. Targeting ultrasound remission in early rheumatoid arthritis: the results of the TaSER study, a randomised clinical trial. *Ann Rheum Dis*. 2016;75:1043–50.
- Haavardsholm EA, Aga AB, Olsen IC, Lillegraven S, Hammer HB, Uhlig T, et al. Ultrasound in management of rheumatoid arthritis: ARCTIC randomised controlled strategy trial. *BMJ*. 2016;354:i4205.
- Rezaei H, Af Klint E, Hammer HB, Terslev L, D'Agostino MA, Kisten Y, Arnaud L. Analysis of correlation and causes for discrepancy between quantitative and semi-quantitative Doppler scores in synovitis in rheumatoid arthritis. *Rheumatology (Oxford)*. 2017;56:255–62.
- D'Agostino MA, Wakefield RJ, Berner-Hammer H, Vittecoq O, Filippou G, Balint P, et al. OMERACT-EULAR-Ultrasound Task Force. Value of ultrasonography as a marker of early response to abatacept in patients with rheumatoid arthritis and an inadequate response to methotrexate: results from the APPRAISE study. *Ann Rheum Dis*. 2016;75:1763–9.
- Backhaus M, Ohrndorf S, Kellner H, Strunk J, Backhaus TM, Hartung W, et al. Evaluation of a novel 7-joint ultrasound score in daily rheumatologic practice: a pilot project. *Arthritis Rheum*. 2009;61:1194–201.
- Aga AB, Berner Hammer H, Christoffer Olsen I, Uhlig T, Kvien TK, van der Heijde D, et al. First step in the development of an ultrasound score for rheumatoid arthritis using data-driven approach. *Ann Rheum Dis*. 2016;75:1444–51.
- Aga AB, Berner Hammer H, Christoffer Olsen I, Uhlig T, Kvien TK, van der Heijde D, et al. Development of a feasible and responsive ultrasound inflammation score for rheumatoid arthritis through a data-driven approach. *RMD Open*. 2016;2:e000325.
- Naredo E, Rodríguez M, Campos C, et al. Validity, reproducibility, and responsiveness of a twelve-joint simplified power Doppler ultrasonographic assessment of joint inflammation in rheumatoid arthritis. *Arthritis Rheum*. 2008;59:515–22.
- Perricone C, Ceccarelli F, Modesti M, Valava C, Di Franco M, Valesini G, et al. The 6-joint ultrasonographic assessment: a valid, sensitive-to-change and

- feasible method for evaluating joint inflammation in RA. *Rheumatology* (Oxford). 2012;51:866–73.
12. Hammer HB, Kvien TK. Comparisons of 7- to 78-joint ultrasonography scores: all different joint combinations show equal response to adalimumab treatment in patients with rheumatoid arthritis. *Arthritis Res Ther*. 2011;13: R78.
 13. Mandl P, Naredo E, Wakefield RJ, Conaghan PG, D'Agostino MA; OMERACT Ultrasound Task Force. A systematic literature review analysis of ultrasound joint count and scoring systems to assess synovitis in rheumatoid arthritis according to the OMERACT filter. *J Rheumatol* 2011;38:2055–2062.
 14. Backhaus TM, Ohrndorf S, Kellner H, Strunk J, Hartung W, Sattler H, Iking-Konert C, Burmester GR, Schmidt WA, Backhaus M. The US7 score is sensitive to change in a large cohort of patients with rheumatoid arthritis over 12 months of therapy. *Ann Rheum Dis*. 2013;72:1163–9.
 15. Ohrndorf S, Halbauer B, Martus P, Reiche B, Backhaus TM, Burmester GR, Backhaus M. Detailed joint region analysis of the 7-joint ultrasound score: evaluation of an arthritis patient cohort over one year. *Int J Rheumatol*. 2013;2013:493848.
 16. Koh JH, Jung SM, Lee JJ, Kang KY, Kwok SK, Park SH et al. Radiographic structural damage is worse in the dominant than the non-dominant hand in individuals with early rheumatoid arthritis. *PLoS One*. 2015;10(8): e0135409.
 17. Möttönen TT. Prediction of erosiveness and rate of development of new erosions in early rheumatoid arthritis. *Ann Rheum Dis*. 1988;47:648–53.
 18. Hammer HB, Bolton-King P, Bakkeheim V, Berg TH, Sundt E, Kongtorp AK, et al. Examination of intra and interrater reliability with a new ultrasonographic reference atlas for scoring of synovitis in patients with rheumatoid arthritis. *Ann Rheum Dis*. 2011;70:1995–8.
 19. Torp-Pedersen ST, Terslev L. Settings and artefacts relevant in colour/power Doppler ultrasound in rheumatology. *Ann Rheum Dis*. 2008;67:143–9.
 20. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;1:307–10.
 21. Tan YK, Moorakonda RB, Allen JC Jr, Chew LC, Thumboo J. Back to the basics: understanding joint swelling and tenderness at the wrist in rheumatoid arthritis through the use of ultrasonography. *Int J Rheum Dis*. 2018; [Epub ahead of print].
 22. Riffbjerg-Madsen S, Christensen AW, Christensen R, Hetland ML, Bliddal H, Kristensen LE, Danneskiold-Samsøe B, Amris K. Pain and pain mechanisms in patients with inflammatory arthritis: a Danish nationwide cross-sectional DANBIO registry survey. *PLoS One*. 2017;12:e0180014.
 23. Lage-Hansen PR, Chrysidis S, Lage-Hansen M, Hougaard A, Ejstrup L, Amris K. Concomitant fibromyalgia in rheumatoid arthritis is associated with the more frequent use of biological therapy: a cross-sectional study. *Scand J Rheumatol*. 2016;45:45–8.
 24. Hammer HB, Michelsen B, Provan SA, Sexton J, Lampa J, Uhlig T, Kvien TK. Tender joint count may not reflect inflammatory activity in established rheumatoid arthritis patients - results from a longitudinal study. *Arthritis Care Res (Hoboken)*. 2018. <https://doi.org/10.1002/acr.23815>.
 25. Naredo E, Möller I, Cruz A, Carmona L, Garrido J. Power Doppler ultrasonographic monitoring of response to anti-tumor necrosis factor therapy in patients with rheumatoid arthritis. *Arthritis Rheum*. 2008;58:2248–56.
 26. Pascual-Ramos V, Contreras-Yañez I, Cabiedes-Contreras J, Rull-Gabayet M, Villa AR, Vázquez-Lamadrid J, et al. Hypervascular synovitis and American College of Rheumatology Classification Criteria as predictors of radiographic damage in early rheumatoid arthritis. *Ultrasound Q*. 2009;25:31–8.
 27. Brown AK, Conaghan PG, Karim Z, Quinn MA, Ikeda K, Peterfy CG, Hensor E, Wakefield RJ, O'Connor PJ, Emery P. An explanation for the apparent dissociation between clinical remission and continued structural deterioration in rheumatoid arthritis. *Arthritis Rheum*. 2008;58:2958–67.
 28. Hammer HB, Kvien TK, Terslev L. Ultrasound of the hand is sufficient to detect subclinical inflammation in rheumatoid arthritis remission: a post hoc longitudinal study. *Arthritis Res Ther*. 2017;19:221.
 29. Zufferey P, Brulhart L, Tamborini G, Finckh A, Scherer A, Moller B, Ziswiler HR. Ultrasound evaluation of synovitis in RA: correlation with clinical disease activity and sensitivity to change in an observational cohort study. *Joint Bone Spine*. 2014;81:222–7.
 30. Terslev L, Østergaard M, Sexton J, Hammer HB. Is synovial hypertrophy without Doppler activity sensitive to change? Post-hoc analysis from a rheumatoid arthritis ultrasound study. *Arthritis Res Ther*. 2018;20:224.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

